1989

Recognition, Modularity, and Self

Stephen Arthur
373-1755 Robson Street
Vancouver, BC
Canada V6G 3B7

In free-ranging speculation drawing on arguments by Fodor, Frisby, Hofstadter, Shapiro, and Gazzaniga, I develop a tentative theory whereby cognition in general can be seen entirely in terms of modular recognition and can be equated formally with paranoid delusion. Meaning in general is then seen only as a successful survival mechanism.

Fodor's *Modularity of Mind* (1984) has provided a heuristic prescription for investigating cognitive processes. Modularity refers to the scientifically accessible functions of "sub-cognition" that comprise separable "vertical" faculties, confined to sensory or perceptual domains (e.g., vision or speech recognition), operating swiftly and automatically. What is left is called the central system, processes that operate horizontally across domains on the input provided from the more stereotyped modules (or conversely provide output to equally stereotyped output modules, e.g., speech production). The degree to which input modules are innate or have a fixed neural architecture is irrelevant to the view proposed in this paper. The main point is that the representations transduced by the input modules are akin to conclusions or hypotheses. Recognition is what input modules do, and I will use this idea to construct a particular view of cognition and of that elusive concept, the "self". Where I will differ from Fodor is in my his contention that the split between modules is maintained by isolation from the background belief system. This need not be true if hierarchical, vertical modules tend to intersect at their apexes as competitive schematas. Rather they could appear isolated in the sense that a snowdrift appears stable until triggered into an avalanche (i.e., new belief system). If this is true then most of the work of cognition can be done by the input modules themselves.

In the first section I propose the following paradigm: (a) that perception is really pattern or object *recognition*; (b) that inductive discrimination of prototypes seems to be the basis for the development of the pattern-recognition systems; (c) that this is also what computer neural networks are capable of doing (without central, horizontal rules); (d) that the organization of these neural networks may be similar to the organization of recognition modules; (e) that most of cognition is really subcognition; and so therefore (f) most of cognition could consist entirely of interlocked parallel-distributed-processing (PDP) input modules (including, as input, memory and imagination). A caveat at this point is that this paper is highly speculative, based on broad generalizations and intuitive leaps, and that I do not apologize for any resulting hubris.

Recognition and Cognition

John Frisby in his *Seeing: Illusion, Brain and Mind* (1979) provides a lucid argument that vision is a matter of building up symbols for entities in the scene before us -- that is: recognition. Many visual illusions demonstrate the fact that what is called visual perception really consists of indirect, subjective transformations of the pure sensory input, starting at the lowest level of processing

(in the retina). Visual perception consists of three main stages: (a) A convolution image is produced for each set of ganglion cell receptive field template type; such an image is intrinsically ambiguous. (b) For each cortical minicolumn this set of images is further combined to encode unambiguous features. (c) At higher levels the features are combined in predetermined ways in order to recognize objects. Patterns that constitute object recognition may have developed during postnatal critical periods as what Edelman (1987) calls the secondary repertoire, but nevertheless their inherent bias is akin to the "bug-detectors" of the frog (in Camhi, 1984). While the frog's simple system already encodes object recognition subcortically, the complex human system uses huge areas of cortex extending through the length of the inferior temporal cortex. Separate parallel pathways handle separate subdomains (such as color, motion, space), each operating as a stereotyped, hierarchical input module whose converging output is eventually combined with output of other modules (Changeux, 1985).

It is quite conceivable that there really is a "grandmother neuron". However, such a neuron must exist as part of a distributed network that represents the relatively broad "apex" of the module in question, otherwise lesions would selectively delete single object-recognitions or single memories, and the "graceful decay" found in dementias such as Alzheimer's would not occur (Corteen, Note 1). Analogously, simple computer neural-network recognition modules appear to work by registering schemata by which a whole can be recognized from a part (Hooper and Teresi, 1986). I believe that most of cognition can similarly be described as object recognition of increasing degrees of symbolism. A matching of schematas from converging input/output modules (including perception, memory, imaginations) could occur with "modular" swiftness by a mechanism more akin to natural selection than to computation (Changeux, 1985; Hooper and Teresi, 1986). There is thus little need for any real horizontal, central processor, and perception and cognition could consist of an elaborate form of "priming" (for priming see Solso, 1988).

The phenomenon of cartoons illuminates the mechanism of object recognition (and thereby cognition). A caricature or cartoon can be considered equivalent to Marr's term "primal sketch" in computer vision (in Frisby, 1979), whereby relevant features are grouped and this grouped symbolism is used as the basis for all subsequent analysis. The reason that even children can so readily recognize cartoon-type characters and animals may be due to the fact that cartoons are similar in crucial ways to the kind of structural descriptions of objects that our visual system normally builds up. Thus in Oliver Sacks' description of Dr. P., the man who literally mistook his wife for a hat, Dr. P. is able to recognize objects only to the extent that stereotyped features provide clear clues as to what he is looking at (e.g., the Ace of Spades); he could recognize caricatures but not his own foot. Here, in a modular fashion, the feature-detection system is intact while the next level, object recognition, is obliterated. Thus if the primal sketch was provided for him he could "see". Whereas Sacks describes Dr. P's condition as failing to see the concrete, I would reverse it and say his failing is in seeing the abstract, because the process of constructing objects is one of abstracting.

In the above sense I would like to suggest that normal cognition operates with a frame of reference which is formally equivalent to the pre-tuned attention of paranoid psychosis. Perception employs "top-down" recognition of gestalts at the pre-attentive stage (review see Solso, 1988), and thus it is fundamentally biased. Later I will attempt to show that a paranoid personality is merely a sharpening of normal human object recognition (and by extension, normal human cognition) and of the normal human self.

I have concentrated on vision as an example, but modular recognition covers a large range. Chomsky's "language acquisition device", for example, would be an example of a largely innate vertical input module (although arguments for this being the result of a general horizontal

development can also be made [review see Corteen, Note 1]), and I suggest that its function is primarily one of recognition. Higher cognitive skills like reading or chess-playing probably occur by development of *de-novo* vertical modules (Corteen, Note 1), which emphasizes the selective advantage of doing as much as possible by relatively instantaneous, intuitive, subcognitive processes. A corollary is that those forms of thought that are incomprehensible to most people (such as general relativity), may be so not because of an impoverished central system, but because of the lack of a suitable automatic vertical-input recognition module.

Intuition may be a form of recognition. Computer neural nets designed to recognize the correct conjugation of verbs produce what I would call intuition -- "knowing" without rational rules -- in that they somehow discover hidden regularities of the supposedly irregular verb forms that cannot be spelled out by explicit rules (Corteen, Note 1).

It is now generally accepted that episodic memory is inaccurate and is highly modifiable by the act of recall and rehearsal itself (see Solso, 1988). This modification is consistent with the idea that memory percepts are a form of recognition -- a re-matching between networks invoked by the original perception -- with the same attendant bias and distortion of an external recognition. In that sense all of the sensory cortex and most of the post-central association cortex and hippocampus is doing one thing: recognizing.

Fodor (1983) speculates that input modules evolved to free the central systems, but one could well argue the contrary: that central systems are a more recent, refined emergent property of the more primary, intersecting input modules. In fact I would expect the mind of a seasnail to be closer to a set of vertical input/output modules akin to a computer neural net than to any sort of real horizontal central system. This suggests that most of the evolutionary expansion of the neocortex is due to an increasing ability to combine recognized percepts into higher recognition objects (symbols). The evolution of the neocortex is thus the evolution of recognition, and the corresponding motor and predictive capacities to match; humans are the supreme recognizers.

My argument (and Fodor's) focuses on input modules, but I would suggest that the same reasoning may apply to output modules "in reverse", for example, in the way that the recognition schemata that constitute declarative memory can become transformed into procedural memory (an automatic modular function) (Corteen, Note 1).

The problems that the classic models of selective attention face in attempting to discriminate between mechanisms of unconscious, preconscious and conscious perception or memory may be due to the difficulty of pinning down central processing functions (Treisman; Deutsch and Norman; Craik and Lockhart; in Solso, 1988). On the other hand, the fact that they all tend to adopt a horizontal-faculty approach may be the source of the problem.

Recognition is in a way "arbitrary" and "autistic". Sacks' encounter with the autistic idiot-savant twins (1987) who communicated with each other by way of prime numbers is an example of perception as alien as a bug-detector yet in some narrow realm profoundly perceptive. Modularity is indicated by Sacks' observations of cases where functionally circumscribed faculties can "kick in or out" in an otherwise intact brain: loss of speech recognition; loss of object recognition of one's own leg or hands; loss of proprioception; loss of recognition of place; even loss of recognition of the left half of the world (this woman could not even turn to the left because the concept was unimaginable to her); and on the other hand "additions" such as phantom limbs and the desperate confabulations of a severe Korsikov's patient. These cases emphasize the arbitrariness with which a recognition process can be narrowed, "appended" or "lopped off" independently from all other systems, and in fact without the person recognizing the sudden loss of recognition. Other non-organic, non-psychotic experiences reveal the same kind of phenomena -- deja vu, jamai vu, depersonalization, derealization,

misidentification, sense of presence, autoscopy, etc. (Reed, 1988). The arbitrary nature of recognition processes is emphasized here in order to support the following speculations about the close similarities between the normal and the deluded "self", and the nature of "meaning" in general.

Self

One might use the word ego to describe the "puppeteer" activity of the brainstem reticular activating system as it directs attention and motivation, but that is not the kind of self I have in mind. I mean the self that is aware of being a self, that process that is the awareness of being a unitary, separate, unique, and continuous space-time pattern.

Hofstadter (1981) describes in exquisite, precise, poetic detail a theoretical model of the brain by which the self may well be the result of recognition turned back on itself. In attempting to separate-out free will from reductionist determinism, Hofstadter develops a model of the brain as an elaborate pinball machine seen at different timescales whereby the view of the mind as an epiphenomenon of molecular determinism, and the converse view of mind symbols "willing" the actions of molecules, can both be seen to be true. The responses of such a system are akin to the responses of a whole country -- even though there is no real continuous, same, unique *entity* from one time to the next, nevertheless there is some recognizable consistency. This incredibly complex, yet entirely mechanical system is designed to recognize and to respond, and one of the things it recognizes is its own responses over time. That recognition then affects the system's responses in a cascading effect, whereby the more that illusion of unity is cycled through the system the more established and locked-in the whole illusion becomes until "it has so deeply permeated the system's structure that no one could possibly explain how or why the system works as it does without referring to its 'silly, self-deluding' belief in itself *as a self*." (p. 626).

So: if one's self is created by an act of recognition, and if recognition is the work of automatic modules, then our sense of self (rather, our *creation* of self) may be unrealistically rigid. It may also be unrealistically arbitrary -- an essential illusion. A supposedly central system such as the self is defined by the action of vertical, stereotyped input modules, and may actually be a module "grown" like a highly-developed skill. In fact, foundlings discovered to have grown up without any human contact appear to lack a sense of self entirely, so that self can be thought of as a social construct (Hooper and Teresi, 1986).

Paranoia

Shapiro (1965) describes the cognitive style of a paranoid personality as follows: Paranoids pay attention not to the apparent facts, but to any feature that tends to confirm their biased expectation (suspicion). Their attention is a relentless scanning and searching; nothing pertaining to their concerns escapes it. They avoid surprise by virtually anticipating it. The resulting projections are a compromise with reality. Paranoia, then, sounds to me very much like the picture I have attempted to draw here of what recognition modules are designed to do, and even how the self is constructed -- only more so. A sense of proportion is lost.

A young boy shakes his bible and damns the blasphemers around him in his elementary school. Across the globe another young boy acts as a human mine-sweeper to gladly serve the cause of the Islamic holy war. These phenomena may occur so frequently precisely because the sophisticated human survival mechanism of recognition is so akin to paranoia. Established religions, in this sense, whether fundamentalist or sophisticated, are indistinguishable from paranoid delusion. Thus highly-developed recognition can be counter productive. Survival mechanisms are often inappropriate in contexts to which they are not designed, so the delusional nature of human nature should not be surprising -- it is a problem

all animals encounter when they wander outside of their niches.

On the other hand, in the same way that paranoid delusion is a compensating mechanism for an individual, arbitrary social codes are a similar and perhaps necessary mechanism for society (Wilson 1978). Koestler (1967) identifies the general human tendency toward fanaticism as the result of a necessary self-transcending side of our nature that is essentially the *recognition* of belonging to a larger social or ideological "self".

Meaning

From his pioneering studies of split-brain patients, Gazzaniga (1985) has developed an interpretation of brain modularity and self that parallels' Hofstadter's. In his model, the brain is composed of various loosely-integrated modules, any of which can somewhat autonomously contribute to the action of the whole person. Only one module, the left-hemisphere speech center, is capable of declarative explanation of the reason for any given action. This kind of activity is apparent in forced-choice tests of split-brain subjects, where the actual reason for a choice by the right hemisphere is unknown to the left, such that the left hemisphere instantaneously and automatically fabricates a reason to explain the choice, and the subject really appears to consciously believe it. Thus the self is again created as an interpretation of regularity imposed on a system that is fundamentally *not* regular. In Gazzaniga's view, the "interpreter" is there as an adaptive mechanism to explain or give meaning to phenomena in the world around us. It "turns back onto" the individual in the same manner as Hofstadter's model, but in this case the self is more of a real illusion. Whether "giving of meaning" is tied closely to recognition is perhaps debatable.

The result of Gazzaniga's model is that meaning, itself, can be seen as an evolutionary side-effect. Natural selection has favored a highly-developed and generalized interpretation module. The human condition is thus the state of asking the meaning of meaning, which is tautological in the deepest sense. Thus meaning itself can be seen to be fundamentally "meaningless"; more accurately it is an automatic, modular, systematic survival mechanism, no more, and no less. Even though it is meaningless, however, it is as "true" as any other animal's successful survival mechanism, no more and no less. It is so successful that it applies itself automatically to literally everything, and cannot be stopped. We will be plagued forever with unanswerable questions (what created the universe? if it was a god what created the god?) unanswerable because each answer merely prompts another "why". Thus, again, recognition of meaning (whatever that is) can become counter productive (and thus Frances Farmer needed a lobotomy in order to become a talkshow hostess).

From my own point of view I must admit that this present attempt to explain all of human cognition as forms of recognition is formally equivalent to paranoia. Zen Buddhism is presumably the attempt to go a step beyond all this recognizing.

REFERENCES

Camhi, J. *Neuroethology: Nerve Cells and the Natural Behavior of Animals.* Sunderland, Mass.: Sinauer Ass. 1984

Changeux, J-P. *Neuronal Man.* N.Y.: Oxford, 1985.

Edelman, G. *Neural Darwinism: The Theory of Neuronal Group Selection.* N.Y.: Basic Books, 1987.

Fodor, J. *The Modularity of Mind.* M.I.T.: Bradford Books, 1983.

Frisby, J. *Seeing: Illusion, Brain, and Mind.* Oxford, 1979.

Gazzaniga, M. *The Social Brain.* N.Y.: Basic Books, 1985.

Hofstadter, D. Who Shoves Whom Around Inside the Careenium? In Hofstadter,

D. (Ed.) *Metamagical Themas*. N.Y.: Basic Books, 1985.

Hooper, J. & Teresi, D. *The Three Pound Universe*. N.Y.: Dell, 1986.

Koestler, A. *The Ghost in the Machine*. London, 1967.

Reed, G. *The Psychology of Anomalous Experience*. Buffalo, N.Y.: Prometheus Books, 1988.

Sacks, O. *The Man Who Mistook His Wife for a Hat*. N.Y: Harper and Row, 1987.

Shapiro, D. *Neurotic Styles*. N.Y.: Basic Books, 1965.

Solso, R. *Cognitive Psychology* (Second Edition). Newton, Mass.: Alyn and Bacon, 1988.

Wilson, E.O. *On Human Nature*. Cambridge: Harvard, 1978.

REFERENCE NOTES

1. Corteen, R. Lectures for Psychology 309, "Cognitive Processes", University of British Columbia, Winter Session 1988.